基于深度学习的多种农产品供需预测模型

庄家煜[1,2]，许世卫[1,2*]，李　杨[3]，熊　露[1,2]，刘克宝[3]，钟志平[1,2]

（1. 中国农业科学院农业信息研究所，北京 100081；2. 农业农村部农业信息服务技术重点实验室，北京 100081；3. 黑龙江省农业科学院农业遥感与信息研究所，黑龙江哈尔滨 150086）

摘　要：为进一步提高农产品供需过程模拟与估算精度，本研究以自1980年以来国家级和省级的大量农业数据作为样本，充分考虑农产品品种、时间、收入、经济发展等因素影响，构建基于深度学习长短时记忆神经网络（Long Short-Term Memory Neural Network，LSTM）的多种农产品供需预测模型。模型在充分考虑机理性约束条件的前提下，利用深度学习算法在非线性模型分析预测中的优势，对稻谷、小麦、玉米、大豆、猪肉、禽肉、牛肉、羊肉、水产品等9种主要农产品供需进行分析预测。将基于本模型的2019—2021年产量预测结果与国家统计局公布数据进行对比验证，三年平均预测准确率96.98%，表明本研究构建的预测模型能够高效地反映隐性指标变化对预测结果的影响。该模型可以通过及时地监测农业运行数据，为多区域、跨期的农业展望工作提供智能化技术支持。

关键词：深度学习；供需预测模型；长短时记忆神经网络；循环神经网络；农产品产量；农业展望

中图分类号：F323.7；TP183　　**文献标志码：**A　　**文章编号：**SA202203013

引用格式：庄家煜，许世卫，李杨，熊露，刘克宝，钟志平．基于深度学习的多种农产品供需预测模型[J]. 智慧农业(中英文), 2022, 4(2): 174-182.

ZHUANG Jiayu, XU Shiwei, LI Yang, XIONG Lu, LIU Kebao, ZHONG Zhiping. Supply and demand forecasting model of multi-agricultural products based on deep learning[J]. Smart Agriculture, 2022, 4(2): 174-182. (in Chinese with English abstract)

1　引　言

“民为国基，谷为民命”，粮食安全事关国计民生，监测预警是现代农业管理的高端工具和保障国家粮食安全的重要手段。农业展望作为监测预警体系建设的重要工作，既能增强政府对农产品市场调控的主动权，又能提升农民、商户等各类市场主体应对市场变化的掌控权[1]。然而，农产品品种多、产业影响因素杂、品种间关联性强，无疑对农产品生产、消费、贸易等主要环节实时精准量化分析预警工作提出挑战[2]。

农产品供需分析预测有助于积极释放市场信号、引导农业生产与消费，一直受到国内外研究者的广泛关注，其研究大体可分为农产品产量预测和需求预测两方面。其中，产量预测研究多基于气象[3,4]、遥感[5,6]和作物生长机理[7,8]等模型，结合线性回归、时间序列[9]、人工智能[10]等方法进行预测；需求预测研究多采用人均营养摄取推算法[11,12]、趋势和经验估算法[13,14]、结构

收稿日期：2022-03-23

基金项目：中国农业科学院创新工程项目（CAAS-ASTIP-2016-AII）；中国农业科学院人才项目（JBYW-AII-2022-04）；国家自然科学基金项目（21974012）

作者简介：庄家煜（1982—），男，博士，副研究员，研究方向为农业信息分析。E-mail：zhuangjiayu@caas.cn

*通信作者：许世卫（1962—），男，博士，研究员，研究方向为农业信息分析。E-mail：xushiwei@caas.cn

模型预测法[15, 16]等。农业系统具有很强的非线性、随机性和动态特征[17]，其模型的有效应用依赖于参数的快速、准确估算。许多研究者认为智能优化算法应用到作物模型参数估算中，如利用个体优势遗传算法优化水稻生育期模型参数，实现模型参数的自动率定，可有效提高模型的调参效率[18, 19]。王盈旭等[20]证明通过改进的深度学习算法可以适合非线性系统的建模，并能得到很好的预测效果。Deléglis等[21]发现深度学习算法适合处理粮食供需相关多源异构数据且预测效果较好。Emerson等[22]、Schwalbert等[23]及Shahhosseini等[24]进一步研究表明，应用人工神经网络（Artificial Neural Network，ANN）等智能学习方法预测农作物产量较传统时间序列和多元线性回归模型效果要好。

智能化模型方法综合分析方法已成为农产品供需分析预测的有效工具。从现有研究来看，传统的统计方法不能全方位对农产品供需过程进行定量描述，需充分发挥深度模型所具备的时间、空间数据连续等优势，进一步提高农产品供需过程模拟与估算精度。本研究结合农学、气象学、经济学及计算机等多学科知识，在撰写《中国农业展望报告》过程中积累的大量数据基础上，充分考虑农产品品种、时间、收入、经济发展等因素影响，应用长短时记忆神经网络（Long Short-Term Memory Neural Network，LSTM）等智能算法在多维变量非线性特征感知中的优势，优化传统农产品供需预测模型，构建多时间维度不同农产品供需预测模型，不仅可以根据某一农产品供需预测模型集群智能调配构建另一种农产品的模型集群，还可以针对同一农产品智能调配建立不同时间尺度的模型集群，提高农产品供需过程模拟与估算精度的同时，实现对农产品供需信息实时精准监测分析预警，并可为多区域、跨预测期的农业展望工作提供有效技术支撑。

2 数据与方法

2.1 数据来源与预处理

2.1.1 数据来源

农产品供需预测模型使用的数据为1980年以来国家级和省级的数据，涉及的指标有：产量、作物种植面积、畜产品出栏量、口粮消费、工业消费、饲用消费、种用消费、损耗和其它消费、进出口、价格、GDP、人口、城镇人口、农村人口、气象要素等，数据来自于官方发布的统计年鉴或报告（《中国统计年鉴》《全国农产品成本收益资料汇编》、海关总署进出口数据库等）。部分数据通过调查、回溯等方法进行估算，如饲用消费、工业消费、种用消费和其它消费。

具体的，作物生产相关影响因素包括：降雨量、日照时数、温度、成本收益情况、费用和用工情况、化肥种子投入、支持政策、保护政策和科技政策；畜禽生产相关影响因素包括：出栏量、存栏量、育龄结构、成本收益情况和管理因素；农产品消费相关影响因素包括：口粮消费、工业消费、饲用消费、种用消费、损耗、人均收入、市场价格、人均国民生产总值、工业增长率、畜产品产量、料肉比、饲料价格、投入品和产出品的价格、每亩用种量（作物类）和损耗系数等。

2.1.2 数据预处理

Bengio等[25]发现导致双向循环神经网络（Recurrent Neural Network，RNN）训练极其困难的根本因素在于梯度消失（或梯度爆炸），因此在使用模型计算之前，需要对数据进行预处理。由于本研究涉及的农产品供需预测模型数据来源众多、数据量庞大，为消除指标之间的量纲影响，对不同指标所具有不同的量纲进行数据标准化处理，以解决数据指标之间的可比性问题。对不完整的数据集进行处理，以保证分析结果的可靠性。分别采取均值、热卡填充、k最近邻、回归分析等方法进行填充，采用划分比例、趋势分析等方法对数据进行拆分或合并。对异常数据

进行判定和处理，主要包括拉依达准则（3δ）、肖维勒准则（Chauvenet）、狄克逊准则（Dixon）、罗马诺夫斯基准则（t检验）和格拉布斯准则（Grubbs）等。模型所需数据的预处理包括农产品供需数据的质量分析、异常检测、填充、转换、标准化和归一化等。其中数据的转换和归一化在模型内部通过参数的形式根据模型需要自主选择，其他数据处理方式通过数据导入过程实现。数据转换的方式包括：

（1）求解差分数据。当前时间点的数据与前一时间点的数据进行差分处理，通过差分后的值进行预测。

（2）数据对数处理。对于不符合正态分布的数据计算偏度，取对数使数据近似服从正态分布。

为使LSTM深度学习模型尽快达到收敛，模型中对所有特征数据采用最大最小值归一化方法，归一化之后的数据直接用于建模预测，最后预测结果根据归一化过程中记录的最大或最小值换算还原成实际值。

2.2 研究方法

本文涉及的农产品供需预测模型主要考虑中国农产品供给与需求的特点，并基于均衡原理假设：农产品的供给和需求共同决定了价格，且农产品的总供给为国内总产量、进口量、期初库存之和，总需求为国内消费量、出口量、期末库存之和。在本模型构建时，为深入分析主要农产品的消费需求，对中国国内消费量进行了更详细的划分，根据消费用途将粮食类消费分成口粮消费、饲用消费、工业消费、种用消费、损耗和其它消费；其中，大豆分为食用消费、压榨消费、种用消费、损耗和其它消费。不同用途的粮食消费所受到的影响因素也不相同。口粮（食用）消费受人口和城镇化率水平影响；饲用消费主要受饲养所需的农产品产量影响，工业用粮主要受工业产品产量影响，种用消费主要取决于亩均用种量和播种面积；损耗和其它消费主要取决于产量和损耗系数。

本研究提出的中国农产品供需预测模型主要有两个特点：一是系统考虑农业与非农业、农业内部之间的各种产业联系，包括消费和生产方面的联系，不同产品之间的相互影响，这些产品的产出共同构成GDP，并决定居民收入，进而决定居民消费需求；二是将供需均衡基本假设作为先验知识融合到模型中，并作为约束条件，再引入LSTM方法作为智能预测模型的核心算法。

2.2.1 LSTM算法

本研究设计的深度神经网络结构的关键部位包括InputGate（输入门），OutputGate（输出门）和ForgetGate（遗忘门），这些Gate（门）的实现由Sigmoid函数输出，且取值都在0～1之间，可与节点h（.）进行逐点相乘，以实现“门”的功能。

其中输入门、输出门和遗忘门的更新方法如公式（1）～公式（3）所示。

$$\begin{cases} in_t = \sigma(W_{in}[h_{t-1}, x_t] + b_{in}) \\ C_t = \tanh(W_c[h_{t-1}, x_t] + b_c) \\ C_t = f_t \cdot C_{t-1} + in_t \cdot C_t \end{cases} \tag{1}$$

$$\begin{cases} out_t = \sigma(W_i[h_{t-1}, x_t] + b_{out}) \\ h_t = out_t \cdot \tanh(C_t) \end{cases} \tag{2}$$

$$f_t = \sigma(W_f[h_{t-1}, x_t] + b_f) \tag{3}$$

其中，in_t，out_t和f_t分别是输入门、输出门和遗忘门；σ是遗忘因子；C_t是当前层LSTM的输出；C_{t-1}是LSTM上一层的输出；x_t是当前层的输入；h_{t-1}是上一层到当前层的输入；h_t是当前层到下一层的输出。

输出与下一个输入和隐藏状态有关，本研究将使用RNN来计算该关系，时间步长t的前向隐藏状态通过以下公式（4）计算：

$$h_t^f = \sigma(W_h^f[h_{t-1}, x_t] + b_h^f) \tag{4}$$

其中，h_t^f为向前隐藏状态；b_h为偏置。

利用公式（5）计算后向隐藏状态：

$$h_t^b = \sigma(W_h^b[h_{t-1}, x_t] + b_h^b) \tag{5}$$

最后，利用公式（6）计算时间步长t处的输出：

$$out_t = \left[h_t^f, h_t^b\right]W_{h,out} + b_{out} \quad (6)$$

2.2.2　模型建立

农产品供需预测模型涵盖稻谷、小麦、玉米、大豆、油料、棉花、糖料、马铃薯、蔬菜、水果、猪肉、禽肉、牛肉、羊肉、禽蛋、奶制品、水产品和工业饲料共计18种农产品的供需平衡表和价格预测。在多品种模型运算时，粮食类农产品的饲用消费需要建立在畜禽产品的生产基础之上，所以建模时将猪肉（PK）、牛肉（BV）、羊肉（MU）和禽肉（PT）等肉类品种划分为优先计算组，通过模型计算出肉类联合平衡表方程，再将预测的肉类产量（视结果可用性，若不可用则以存栏量、出栏量等变量替代）作为输入数据引入稻谷（RI）、小麦（WT）、玉米（MA）和大豆（SB）等主要粮食作物品种的联合平衡表方程，计算有种植竞争、消费替代的关联农产品。

供需预测模型总体思路为：将深度神经网络训练模型封装为一个I/O（输入/输出）开放的模块化模型，预留出外部数据输入控制接口，实现供需平衡相关指标预测和平衡表的换算。主要功能包括模型自动训练封装、模型自动调参、模型接口参数输入、平衡表的自动换算、自动配平和最终平衡表输出等。上述方案可以实现农作物、肉类和禽蛋类等各种类别平衡表的自动平衡和品种间关联耦合，并可以动态测算未来10年内的状况。模型可以通过选择待预测的农产品品种、模型训练的历史数据时间段，也可选择未来的几年的平衡表的输出（中长期预测），对模型预测场景进行设置。

模型的输入包括农产品自2000年以来预测的平衡表数据，以及平衡表数据（包括生产量、消费量、进口量、出口量、结余变化）、生产量和消费量的组成分项数据（包括单产、面积、亩均用种量、自损系数、损耗系数、城镇人口比例、出栏量、胴体重等）、年度价格数据、宏观经济数据（人口、CPI等价格指数、GDP、工业指数等）、国际货币数据（人民币与主要货币的汇率）等。模型的输出为各类农产品自预测时间以来未来十年的平衡表。平衡表测算模型主要采用LSTM模型方法测算，辅以逻辑回归（Logistic Regression，LR）模型进行局部变量的配平。在LSTM中，测算主要是通过多输入多输出，系统自动建模得出。其中，所有的测算都可以设置限制条件，如设定输入变量因素的权重等。参考经济学中均衡理论，供给和需求影响价格，同时被价格影响，因此建立生产量、消费量、进出口量、结余变化、价格和受价格影响的宏观变量及指标的多输入、多输出模型。

农产品供需平衡模型的基本约束方程如公式（7）～公式（12）：

总供给 = 总需求　（7）

总供给 = 生产量 ＋ 进口量 ＋ 期初库存　（8）

总需求 = 消费量 ＋ 出口量 ＋ 期末库存　（9）

消费量 =口粮消费量 ＋ 工业消费量 ＋ 种用消费量 ＋ 损耗［ ＋ 饲用消费量］　（10）

当期期初库存 = 上一期期末库存　（11）

生产量 = 单产 ×播种面积　（12）

其中，“［ ］”中的内容表示可选项。

2.2.3　模型数据接口

供需预测模型的输入为每组包含 m 个特征的 n 组历史数据，输出为每组包含 m 个特征的 k 组预测数据。通常，模型输出数据要求越多，预测结果的精度也会相应降低。因此，本研究设计的LSTM模型输入输出数据满足如下形式：

$(\boldsymbol{X},\boldsymbol{Y})$=[样本,时间步长,特征]　（13）

其中，$\boldsymbol{X}$ 为影响农产品供需的各类影响因素向量，各因素之间需要根据其含义进行特定的特征衍生；$\boldsymbol{Y}$ 为供需平衡表各项形成的向量，所有特征在进入网络前需要进行标准化和归一化处理，同时需要控制参数对 $\boldsymbol{Y}$ 分项进行边界限制。对于LSTM的多输入多步长多输出数据，使用Encoder-Decoder（编码-解码模型）LSTM，其对于多输入多步长多输出的序列到序列（Sequence to Sequence）问题具有比较好的效果。

2.2.4　模型评估方法

通常，预测模型平均损失越小越好，且损失

曲线趋于收敛。在模型集群验证时，将历史数据进行滑动处理，以最新的一年数据作为验证集，之前的数据用于建模训练，将真实值和模型预测值做对比验证。例如，为验证2015年预测结果，则将1980—2014年数据作为训练集，2015年实际数据作为验证集；同理，验证2016年预测结果，则将1980—2015年数据作为训练集，2016年实际数据作为验证集，以此进行滑动验证。

本研究涉及的模型评估主要使用平均绝对值误差（Mean Absolute Deviation，MAE）和平均绝对百分比误差（Mean Absolute Percentage Error，MAPE）两个指标判断模型的效果。

3 结果与分析

3.1 单品种建模预测

以耦合度较小的禽蛋供需预测（2021—2030年）为例，预测时使用双输入模型，将样本数据分为禽蛋历史平衡表数据和影响禽蛋供需的各要素数据。其中，卷积层输入数据为影响禽蛋供给与需求各分项的影响因素数据，卷积处理后对数据进行特征提取。在全连接层接收平衡表分项数据，与先前的因素特征组合成新的特征，通过LSTM算法进行禽蛋供需预测，流程如图1所示。

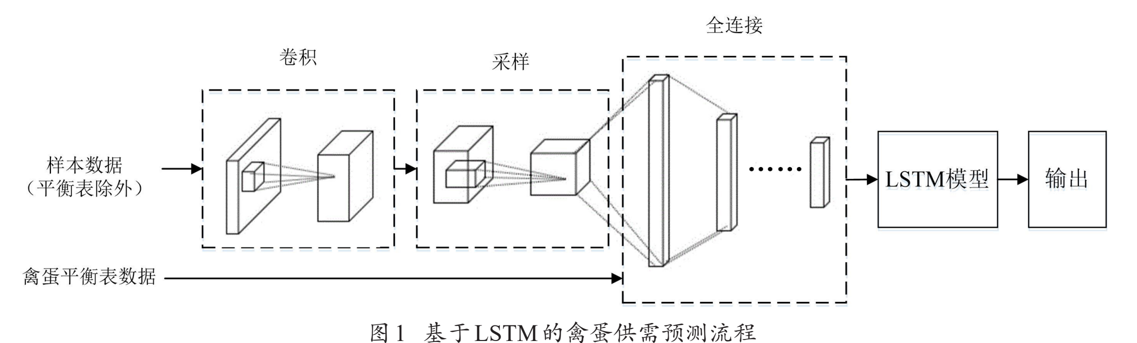


图1 基于LSTM的禽蛋供需预测流程

Fig. 1 Egg supply and demand forecasting process based on LSTM

图1中模型网络采用串联结构，可以提高模型的表达能力。模型采用模块分组的形式，以模块为基础，重复提取不同数量的特征。模块的基本构成为：先对输入进行3×3的卷积，输出为n/4个通道；然后连接2×2的卷积，输出为n/4个通道；对数据进行两次特征变换，最后加入一个1×1的卷积，提高网络的表达能力，同时组合不同通道的信息，输出为n个通道（其中n为滤波器的个数）。在主体网络结构中引入模型池化层，以合成初始特征区域内某一个特征，并有效减小过拟合现象的发生。丢弃层（Dropout）使隐层神经元随机失活，防止模型过拟合，同时可以提高模型的泛化能力。由于样本数据存在分布不均匀现象，采用加权交叉熵作为目标的损失函数。全连接层加入时序信息特征，综合考虑时间特征。对于预测出的结果，需要满足平衡表的分项之间的关系，同时使用Lasso回归模型进行校正。首先，将禽蛋的历史数据处理成一系列的时间序列，并利用LSTM模型学习其周期特征，训练过程中损失随训练次数（Epoch）的变化曲线如图2所示。根据中间过程的输出信息，在迭代过程中损失函数在减小。

模型预测的未来10年中国禽蛋供需结果如表1所示，表中2011—2020年为中国禽蛋历史供需平衡表，2021—2030年为模型预测值。

3.2 多品种集群建模预测

与上述禽蛋供需预测方法类似，通过LSTM方法优化的农产品供需预测模型可以实现多品种集群建模预测。在产量预测方面，主要使用气象数据、投入和管理政策等影响因素；在消费预测方面，建立差异化消费预测，引入人口、农产品价格、地区消费差异等因素，为不同消费类型的消费量计算提供定量分析技术；在价格预测方

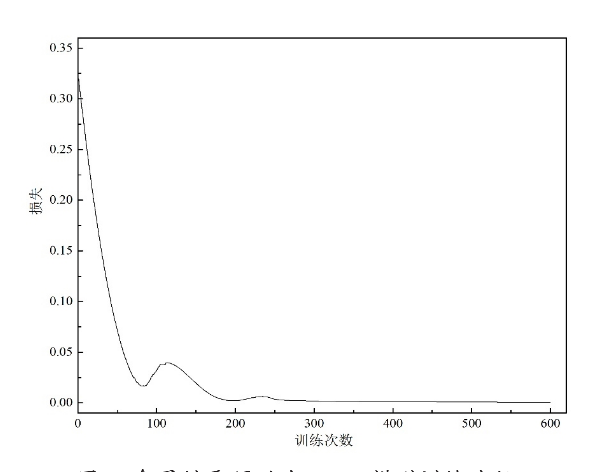


图 2 禽蛋供需预测的 LSTM 模型训练过程

Fig. 2 Training process of egg supply and demand forecast LSTM model

面，引入多品种价格关联和竞争机制，建立多维农产品价格预测方法。

利用本研究供需预测模型对中国主要农产品（稻谷、小麦、玉米、大豆、猪肉、禽肉、牛肉、羊肉和水产品）产量进行预测，预测结果与国家统计公布的实际结果进行对比，结果如图 3 所示。产量预测模型的三年（2019—2021 年）总体 MAPE 为 3.02%，即三年平均预测准确率为 96.98%。其中 2019 年 MAPE 为 3.90%，2020 年 MAPE 为 1.74%，2021 年为 3.42%，即准确率分别为 96.10%、98.26% 和 96.58%。可以看出 2020 年和 2021 年的产量预测精度要好于 2019 年，说明随着样本容量的增加，智能学习模型预测效果会逐步变好。

表 1 禽蛋历史与预测平衡表

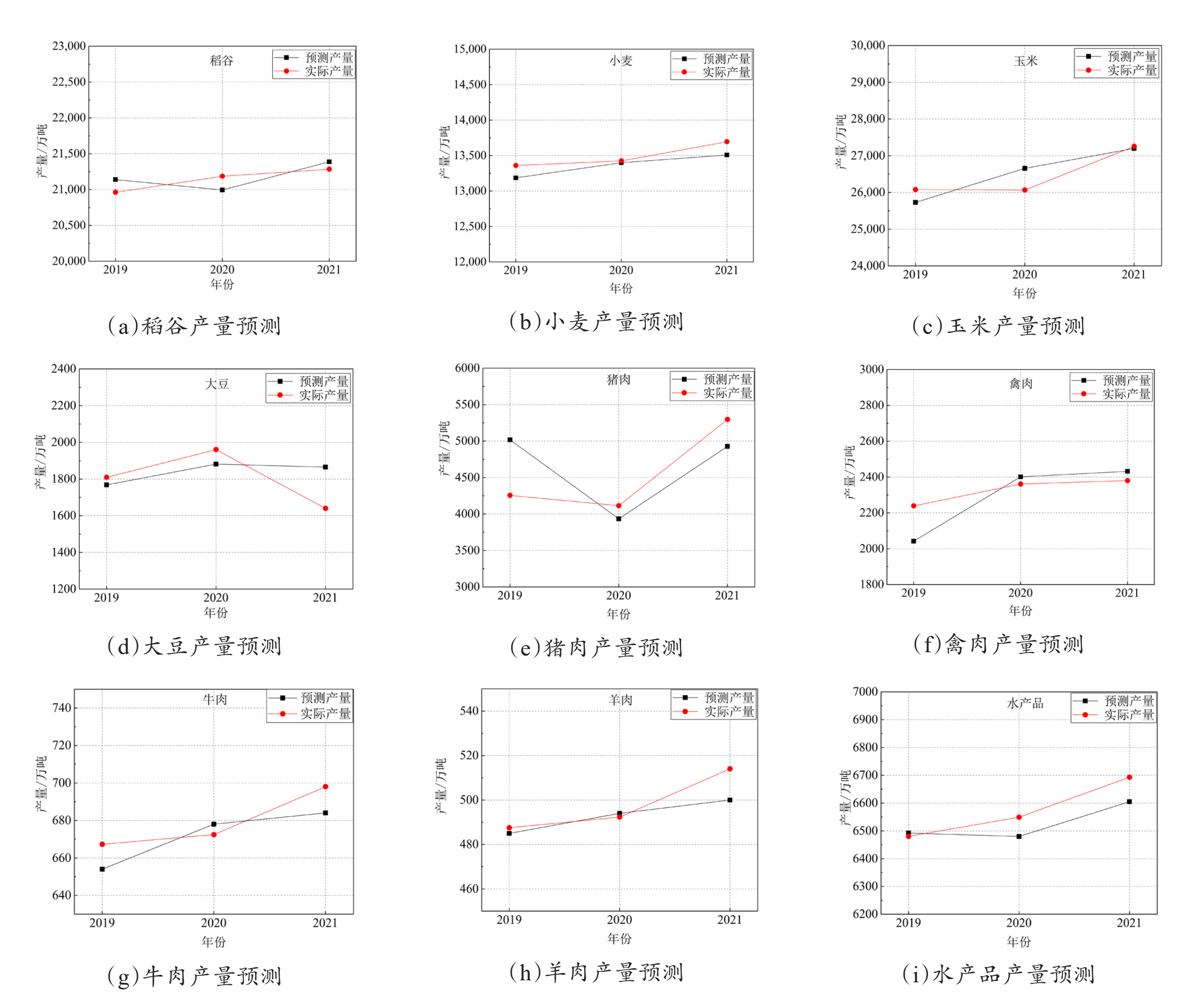
Table 1 History and forecast balance sheet of egg

单位：万吨

	年份	产量	进口量	消费量	食用消费	加工消费	种用及损耗	出口量	结余变化
历史数据	2011	2830	0.007	2815	2072	496	248	10	5
	2012	2885	0.003	2867	2119	498	250	10	8
	2013	2906	0.002	2918	2167	500	251	9	−22
	2014	2930	0.002	2973	2219	502	253	9	−52
	2015	3046	0.002	3032	2326	458	248	10	4
	2016	3161	0.000	3145	2427	467	252	10	5
	2017	3096	0.006	3094	2369	476	249	11	−9
	2018	3128	0.000	3120	2386	483	251	10	−2
	2019	3309	0.002	3296	2542	491	263	10	3
	2020	3468	0.013	3449	2662	516	271	10	9
预测数据	2021	3406	0.003	3393	2598	526	269	11	2
	2022	3449	0.003	3438	2623	541	274	11	0
	2023	3488	0.003	3475	2644	552	279	11	2
	2024	3514	0.003	3504	2662	561	282	11	−1
	2025	3551	0.003	3535	2679	569	287	11	5
	2026	3577	0.003	3554	2689	576	288	11	12
	2027	3592	0.003	3573	2701	583	290	12	7
	2028	3608	0.002	3592	2710	590	292	12	4
	2029	3631	0.002	3612	2719	598	294	12	7
	2030	3650	0.002	3629	2726	608	295	12	9

4 总结与展望

由于传统农产品供需统计分析方法不能全方位对生产、流通、消费的全过程进行定量化描述，为充分发挥大数据应用的时间连续、空间数据连续等优势，进一步提高多品种供需的模拟与估算精度，本研究提出了一种基于深度学习模型



(a)稻谷产量预测　(b)小麦产量预测　(c)玉米产量预测

(d)大豆产量预测　(e)猪肉产量预测　(f)禽肉产量预测

(g)牛肉产量预测　(h)羊肉产量预测　(i)水产品产量预测

图3　基于LSTM模型的2019—2021年主要农产品产量预测结果与实际对比

Fig. 3　2019-2021 production forecast and actual results of major agricultural products based on the LSTM model

LSTM的农产品供需分析预测模型，可实现多元高频数据协同分析，有效解决农产品供需分析过程中因单一信息源造成的“视觉盲区”，对全息精准进行农产品供需预测预警具有一定的探索性和开创性。

本方法在对农产品供需模型参数敏感性分析和参数校正基础上，以1980年以来中国农产品供需平衡表数据作为基础学习样本，构建了基于长短时记忆神经网络和均衡关联的农产品供需预测模型。可实现稻谷、小麦、玉米、大豆和禽蛋等主要农产品供给、需求的定量化模拟与估算。本研究构建的农产品供需预测模型，利用均衡理论对农产品供需各个分量进行约束，同时综合考虑农产品产业链各环节的历史特征以及影响因素，避免了人工经验干预引入的不可控误差，并通过循环神经网络优化了模型的参数，提高了模型的预测性能。

农产品供需预测模型是中国农业展望工作的核心，该类模型的应用可以提升“用数据说话、用数据决策、用数据管理、用数据创新”的技术支撑能力，为农业生产经营和宏观调控提供市场预警、政策评估、资源管理和舆情分析等数据产品与技术支持。本研究所构建模型尚存在一些不足，对诸如非洲猪瘟、中美贸易摩擦等特殊事件的响应还不够。这是因为重大政策、突发事件等外部冲击因素属于偶发性干扰变量，很难获取大量的样本特征，并通过深度学习方法形成相应的智能模型。所以，未来农产品供需预测模型还需

要和更多的模型进行协同应用，才能更好地对农产品供需形势进行定量化分析预测。

参考文献：

[1] 农业农村部市场预警专家委员会. 中国农业展望报告(2021—2030)[M]. 北京: 中国农业科学技术出版社, 2021.

[2] 许世卫, 邸佳颖, 李干琼, 等. 农产品监测预警模型集群构建理论方法与应用[J]. 中国农业科学, 2020, 53(14): 2859-2871.
XU S, DI J, Li G, et al. The methodology and application of agricultural monitoring and early warning model cluster[J]. Scientia Agricultura Sinica, 2020, 53(14): 2859-2871.

[3] ZHUANG J, XU S, LI G, et al. The influence of meteorological factors on wheat and rice yields in China[J]. Crop Science, 2018, 58(3): 837-852.

[4] LAUDIEN R, SCHAUBERGER B, MAKOWSKI D, et al. Robustly forecasting maize yields in Tanzania based on climatic predictors[J]. Scientific Reports, 2020, 10: ID 19650.

[5] JOHNSON M D, HSIEH W W, CANNON A J, et al. Crop yield forecasting on the Canadian Prairies by remotely sensed vegetation indices and machine learning methods[J]. Agricultural and Forest Meteorology, 2016, 218-219: 74-84.

[6] EUGENIO F C E, GROHS M, VENANCIO L P, et al. Estimation of soybean yield from machine learning techniques and multispectral RPAS imagery[J]. Remote Sensing Applications, 2020: ID 100397.

[7] PAUDEL D, BOOGAARD H, DE WIT A, et al. Machine learning for large-scale crop yield forecasting[J]. Agricultural Systems, 2020, 187: ID 103016.

[8] DE WITA, BOOGAARD H, FUMAGALLI D, et al. 25 years of the WOFOST cropping systems model[J]. Agricultural Systems, 2019, 168: 154-167.

[9] MASUDA, TADAYOSHI, GOLDSMITH, et al. World soybean production: Area harvested, yield, and long-term projections[J]. International Food & Agribusiness Management Review, 2009, 12(4): 143-161.

[10] 王桂芝, 胡慧, 陈纪波, 等. 基于BP滤波的Fourier模型在粮食产量预测中的应用[J]. 中国农业气象, 2015, 36(4): 472-478.
WANG G, HU H, CHEN J, et al. Application of Fourier model based on BP filter in crops yield prediction[J]. Chinese Journal of Agrometeorology, 2015, 36(4): 472-478.

[11] 肖玉, 成升魁, 谢高地, 等. 我国主要粮食品种供给与消费平衡分析[J]. 自然资源学报, 2017, 32(6): 927-936.
XIAO Y, CHENG S, XIE G, et al. The balance between supply and consumption of the main types of grain in China[J]. Journal of Natural Resources, 2017, 32(6): 927-936.

[12] 谢高地, 成升魁, 肖玉, 等. 新时期中国粮食供需平衡态势及粮食安全观的重构[J]. 自然资源学报, 2017, 32(6): 895-903.
XIE G, CHENG S, XIAO Y, et al. The balance between grain supply and demand and the reconstruction of China's food security strategy in the new period[J]. Journal of Natural Resources, 2017, 32(6): 895-903.

[13] 赵萱, 邵一珊. 我国粮食供需的分析与预测[J]. 农业现代化研究, 2014, 35(3): 277-280.
ZHAO X, SHAO Y. Analysis and forecast of China's grain supply and demand[J]. Research of Agricultural Modernization, 2014, 35(3): 277-280.

[14] 刘洋, 罗其友, 周振亚, 等. 我国主要农产品供需分析与预测[J]. 中国工程科学, 2018, 20(5): 120-127.
LIU Y, LUO Q, ZHOU Z, et al. Analysis and prediction of the supply and demand of China's major agricultural products[J]. Strategic Study of CAE, 2018, 20(5): 120-127.

[15] LU W, NING L C, WEN X Q. Modeling the effects of urbanization on grain production and consumption in China[J]. Journal of Integrative Agriculture, 2017, 16(6): 1393-1405.

[16] 黄季焜. 对近期与中长期中国粮食安全的再认识[J]. 农业经济问题, 2021(1): 19-26.
HUANG J. Recognition of recent and mid-long term food security in China[J]. Issues in Agricultural Economy, 2021(1): 19-26.

[17] 陈锡康, 杨翠红. 农业复杂巨系统的特点与全国粮食产量预测研究[J]. 系统工程理论与实践, 2002(6): 108-112.
CHEN X, YANG C. Characteristic of agricultural complex giant system and national grain output prediction[J]. Systems Engineering-Theory & Practice, 2002(6): 108-112.

[18] 许世卫. 农业信息分析学[M]. 北京: 高等教育出版社, 2013.

[19] 高亮之. 农业模型学[M]. 北京: 气象出版社, 2019.

[20] 王盈旭, 韩红桂, 郭民. 一种基于改进型深度学习的非线性建模方法[J]. 信息与控制, 2018, 47(6): 680-686.
WANG Y, HAN H, GUO M. A nonlinear modeling method based on improved deep learning[J]. Information and Control, 2018, 47(6): 680-686.

[21] DELÉGLISE H, INTERDONATO R, BÉGUÉ A, et al. Food security prediction from heterogeneous data combining machine and deep learning methods[J]. Expert Systems with Applications, 2022, 190: ID 116189.

[22] EMERSON R A, DOS REIS J G M, VENDRAMETTO O, et al. Time series prediction with artificial neural networks: An analysis using Brazilian soybean pro-

duction[J]. Agriculture (Basel), 2020, 10(10): ID 475.
[23] SCHWALBERT R A, AMADO T, CORASSA G, et al. Satellite-based soybean yield forecast: Integrating machine learning and weather data for improving crop yield prediction in southern brazil[J]. Agricultural and Forest Meteorology, 2020, 284: ID 107886.
[24] SHAHHOSSEINI M, HU G, HUBER I, et al. Coupling machine learning and crop modeling improves crop yield prediction in the US corn belt[J]. Scientific Reports, Scientific Reports, 2021, 11(1): 1-15.
[25] BENGIO Y, SIMARD P, FRASCONI P. Learning long-term dependencies with gradient descent is difficult[J]. IEEE Transactions on Neural Networks, 1994, 5(2): 157-166.

Supply and Demand Forecasting Model of Multi-Agricultural Products Based on Deep Learning

ZHUANG Jiayu[1,2], XU Shiwei[1,2*], LI Yang[3], XIONG Lu[1,2], LIU Kebao[3], ZHONG Zhiping[1,2]

(1. *Agricultural Information Institute of CAAS*, *Beijing* 100081, *China*; 2. *Key Laboratory of Agricultural Information Service Technology of MOA*, *Beijing* 100081, *China*; 3. *Institute of Agricultural Remote Sensing and Information of HAAS*, *Harbin* 150086, *China*)

Abstract: To further improve the simulation and estimation accuracy of the supply and demand process of agricultural products, a large number of agricultural data at the national and provincial levels since 1980 were used as the basic research sample, including production, planted area, food consumption, industrial consumption, feed consumption, seed consumption, import, export, price, GDP, population, urban population, rural population, weather and so on, by fully considering the impact factors of agricultural products such as varieties, time, income and economic development, a multi-agricultural products supply and demand forecasting model based on long short-term memory neural network (LSTM) was constructed in this study. The general thought of supply and demand forecasting model is packaging deep neural network training model as an I/O-opening modular model, reserving control interface for input of outside data, and realizing the indicators forecasting of supply and demand and matrixing of balance sheet. The input of model included forecasting balance sheet data of agricultural products, annual price data, general economic data, and international currency data since 2000. The output of model was balance sheet data of next decade since forecasting time. Under the premise of fully considering the mechanical constraints, the model used the advantages of deep learning algorithms in nonlinear model analysis and prediction to analyze and predict supply and demand of 9 main types of agricultural products, including rice, wheat, corn, soybean, pork, poultry, beef, mutton, and aquatic products. The production forecast results of 2019—2021 based on this model were compared and verified with the data published by the National Bureau of Statistics, and the mean absolute percentage error was 3.02%, which meant the average forecast accuracy rate of 2019—2021 was 96.98%. The average forecast accuracy rate was 96.10% in 2019, 98.26% in 2020, and 96.58% in 2021, which shows that with the increase of sample size, the prediction effect of intelligent learning model would gradually get better. The forecasting results indicate that the multi-agricultural supply and demand prediction model based on LSTM constructed in this study can effectively reflect the impact of changes in hidden indicators on the prediction results, avoiding the uncontrollable error introduced by manual experience intervention. The model can provide data production and technical support such as market warning, policy evaluation, resource management and public opinion analysis for agricultural production and management and macroeconomic regulation, and can provide intelligent technical support for multi-regional and inter-temporal agricultural outlook work by monitoring agricultural operation data in a timely manner.

Key words: deep learning; supply and demand forecasting model; LSTM; RNN; agricultural production; agricultural outlook